Junior Management Science

www.jums.academy
ISSN: 2942-1861

Experimental Analysis of Gender Bias in Hiring within STEM Fields

Roya Kazimova

Technical University of Munich

Abstract

This study investigates the presence of gender bias in students' hiring evaluations within STEM fields using a survey-based experimental design with systematically varied resumes. Students were randomly assigned to evaluate one of four fictional CVs, differing by candidate gender and qualification strength. Participants rated the applicants on competence, hireability, salary recommendation, and willingness to offer professional support. The results reveal no consistent bias in favor of male candidates; however, strong male applicants received higher salary suggestions, while female candidates were more likely to be offered professional support. These patterns are primarily driven by interactions between candidate gender and resume strength. The effects are most pronounced among participants from Eastern Europe and older participants, suggesting a role of social background in shaping perceptions. The findings highlight how implicit biases may form prior to labor market entry and underscore the need for early educational interventions to promote equitable hiring perceptions in STEM.

Keywords: gender bias; hiring discrimination; survey experiment; resume evaluation

1 Introduction

In recent years, there has been growing attention to diversity and inclusion in the workplace, especially in areas where women are still underrepresented. Many companies, universities, and policymakers are trying to better understand and reduce hidden biases that can affect hiring and promotion decisions (Wilkinson & Barry, 2020). These efforts are especially important in STEM, where gender gaps remain despite efforts to improve equality (Evagorou et al., 2024). One of the key questions in labor economics and gender studies is whether gender bias continues to influence hiring decisions in these male-dominated sectors. These biases not only limit women's access to employment but also affect long-term career growth, salary progression, and leadership opportunities.[1] According to the Gender Social Norms Index by United Nations Development Programme (2023), 85% of the global population has biased views against women, and more than half believes that men are better suited for leadership positions.

While there is strong evidence that gender bias affects real-world hiring decisions, most existing research has focused on the behavior of employers or experienced professionals. For example, Quadlin (2018) conducted a large-scale field experiment and found that women with high academic performance were rated as competent but less likable and hirable than men with moderate academic records. This suggests that even when qualifications are equal or superior, gender stereotypes can significantly shape hiring outcomes. However, much less is known about how university students, who represent the next generation of hiring professionals, perceive and internalize such biases. This gap is particularly relevant in STEM contexts, where common stereotypes (e.g., "math is a male domain") shape both other people's opinions and personal confidence (Correll, 2001).

This study investigates how university students with diverse academic and professional backgrounds perceive gender bias in candidate evaluation. By focusing on students in interdisciplinary programs combining technical and managerial studies, the research captures perspectives that are highly relevant to future decision-making in both STEM and organizational contexts. Most participants have prior work experience, adding practical insight to their evaluations. The study uses a survey-based experimental design in which participants assess fictional job applicants, allowing a systematic investigation of how gender might affect candidate evaluation. This approach shows how students think about gender bias before they enter professional roles and start making hiring decisions.

Preliminary findings from the survey experiment reveal a more complex picture than anticipated. Contrary to expectations, male candidates were not consistently favored and were in some cases rated lower on hireability compared to equally qualified female applicants. While direct gender effects were limited, interaction patterns suggested that strong male candidates were more likely to receive higher salary rec-

[1] McKinsey & Company, *Women in the Workplace 2024: The 10th-Anniversary Report*, 2024, https://www.mckinsey.com/featured-insights/diversity-and-inclusion/women-in-the-workplace.

DOI: https://doi.org/10.5282/jums/v11i3pp598-610

ommendations but were less likely to receive offers of professional support. Additionally, participants' demographic characteristics, particularly age and region of origin, significantly influenced candidate evaluations. These results point to a nuanced understanding of bias, shaped by social norms, cultural background, and prior work experience.

This study contributes to existing research in three ways. First, it examines whether university students recognize gender biases in hiring, filling a gap in research focused on future professionals rather than current employees. Second, it analyzes the role of demographic factors such as gender, field of study, and work experience in shaping students' hiring evaluations. Third, it explores whether students' perceptions of gender bias translate into biased hiring decisions in a controlled experimental setting. By identifying the extent to which students recognize hiring biases and how these biases affect candidate evaluations, this research contributes to developing educational interventions and hiring policies that promote gender equality in STEM fields. Understanding how future professionals perceive and respond to gender biases can inform diversity initiatives, helping to create more equitable hiring practices in STEM and beyond.

The remainder of the paper is structured as follows. Section 2 presents the literature review on gender bias in hiring, experimental methodologies, and perception research. Section 3 describes the methodological approach, including the survey experiment design and data collection. Section 4 outlines the steps of the data analysis. Section 5 presents the empirical results. Finally, Section 6 offers a discussion of the findings and concludes the paper.

2 Literature Review

Research shows that gender bias still affects hiring decisions, even when men and women have the same qualifications. These biases are often subtle and may not be immediately apparent, but they can still influence how candidates are evaluated. In their famous study of orchestra auditions, Goldin and Rouse (2000) demonstrated that "blind" evaluation – when the screen hides the musician's identity – significantly increases the chance that women will be selected. The results of the study show how even minor cues related to gender can influence decision-making and how removing identity-related information can reduce bias. Bias is especially likely to appear in roles that are typically associated with one gender. Employers may not discriminate against women overall, but may be more hesitant to hire them into male-dominated positions, and vice versa for men. This pattern reflects the persistence of gender stereotypes in the labor market (Rich, 2014). When a woman enters a role in a male-dominated field, her competence may be judged not only by her resume but also by stereotypical perceptions of who is "naturally suited" for the role. Over time, even these small patterns of judgment can create unequal starting points in careers. Moreover, they can accumulate and lead to long-term negative consequences for women's careers, such as harsher evaluations, lower starting salaries, or fewer opportunities for advancement (Hardy III et al., 2022).

Despite growing attention to gender equality, women remain underrepresented in STEM fields, especially at higher academic and professional levels. Large-scale studies have shown that women in STEM often face slower career progression. Holman et al. (2018), for example, analyzed over 36 million scientific publications and found that women are less likely to be listed as lead authors. Since being a lead author is important for academic recognition and promotion, the lack of women in these positions shows how gender inequality limits career growth in STEM. Bias in hiring further reinforces inequality. STEM managers may apply different evaluation criteria depending on the candidate's gender. For instance, men in decision-making roles often emphasize long working hours when reviewing female applicants, while giving less weight to core skills. Even small cues, such as including a note about childcare support in a resume, can shift these perceptions and reduce bias (Friedmann & Efrat-Treister, 2023).

Students already form opinions about gender, fairness, and who is the most suitable for certain roles. These beliefs can also shape what they expect from future workplaces. If students believe that bias and discrimination are common, this can discourage them from applying to certain jobs or make them hesitant about entering male-dominated fields like STEM. Previous research has shown that even experienced professionals, such as university professors, tend to rate identical resumes more favorably when they have a male name (Moss-Racusin et al., 2012). If such bias exists among academics, it raises important questions about how early these patterns develop. Studies show that female students are more likely to see gender discrimination as a serious barrier to career success, while male students often believe it is less common or less important. These differences are influenced by personal experiences, education, and exposure to diversity (Sipe et al., 2009). Since today's students will become tomorrow's professionals and decision-makers, it is important to understand their views now.

Discrimination in hiring is often difficult to measure through surveys or self-reported data because people may not be fully aware of their biases or may not feel comfortable admitting them. For this reason, many researchers use experimental methods to study biased behavior in a more controlled and reliable way, like the resume evaluation method. In this design, researchers send out fictional job applications that are identical in qualifications but differ in one key aspect, such as gender or ethnicity. This makes it possible to see the effect of that variable on hiring decisions. Bertrand and Mullainathan (2004), for example, found that job applicants with white-sounding names received significantly more callbacks than those with black-sounding names, even though their resumes were otherwise the same. In another study, Milkman et al. (2015) applied the same method in academia and found that professors responded less often to requests from female and minority students.

When studying hiring discrimination, researchers commonly employ two main types of experimental designs: field experiments and lab-based experiments. The former involve realistic scenarios conducted in natural settings, such as resume studies, where researchers submit fictitious resumes to real job openings. The latter, on the other hand, are conducted in controlled environments or online, allowing researchers to precisely manage variables and conditions. An important distinction exists between these two methodologies. Field experiments effectively capture real-world hiring decisions and are more likely to reveal actual biases because participants are unaware that they are being studied. In contrast, lab and online experiments provide researchers greater control over the experimental conditions but may not fully reflect real-life decisions, as participants typically face no genuine consequences for their choices. As a result, individuals in lab settings might provide answers they believe are socially desirable rather than truthful (Neumark, 2018). Despite this limitation, lab and online experiments remain valuable for understanding perceptions, attitudes, and cognitive processes within a carefully controlled environment.

The aforementioned studies have clarified how gender discrimination operates at different personal development stages, noting, for instance, that girls' interest in leadership declines during adolescence (Alan et al., 2019), and that women are less likely to negotiate salaries early in their careers (Babcock et al., 2003). These early patterns partially explain why fewer women reach leadership positions.[2] However, it remains unclear how early these tendencies form. Specifically, little is known about whether such biases are already present among students before entering the professional workforce. Addressing this gap is crucial, as understanding students' perceptions could help mitigate biases before they become deeply rooted.

3 Methodology

This chapter outlines the methodology used in this study, detailing the research design, survey structure, data collection process, and analytical approach. Section 3.1 presents the research design, explaining the rationale for using an experimental approach and describing the independent and dependent variables. Subsection 3.1.1 provides a detailed account of the experimental setup and procedure, including how candidate resumes were developed and structured. Section 3.2 describes the survey design, outlining the structure of the questionnaire and the key measures used to assess participants' evaluations and perceptions. Section 3.3 explains the data collection process, detailing the recruitment methods, survey distribution, ethical considerations, and response rates. This structure provides a clear and organized overview of the research process, forming the basis for analyzing gender bias in hiring.

3.1 Research Design

This study consists of two distinct methodological components: experimental survey design and descriptive survey analysis. The first part of the study employs a 2×2 between-subjects factorial experimental design to examine gender bias in hiring evaluations. Participants were randomly assigned to evaluate one of four candidate resumes, varying in gender (male vs. female) and qualification strength (strong vs. weak). This design ensured that each participant assessed only one candidate, preventing potential biases that could arise from direct comparisons. The design simulates a realistic hiring scenario, measuring hireability, competence, salary expectations, and professional support willingness. This method aligns with previous research on labor market discrimination and provides a structured approach to testing whether gender bias influences candidate evaluations. The study manipulates two independent variables: candidate gender and candidate qualification strength. The dependent variables include hireability (likelihood of hiring the candidate), competence (perceived ability and skills), and professional support (willingness to offer help in academic or professional settings). In addition, five demographic characteristics (participant age, gender, country of origin, field of study, and work experience) were included as control variables to account for their potential influence on candidate evaluations.

The second part of the study employs a descriptive survey analysis to examine participants' perceptions of gender discrimination in hiring, professional mentoring, and promotion opportunities. Unlike the experimental section, this part does not manipulate variables but instead presents a summary of responses, calculating mean discrimination perception scores across different participant groups (e.g., age, gender, country of origin, educational background, and work experience). Additionally, three control variables such as discrimination against women, men, and oneself, were derived from participants' responses. A total of 92 completed responses were included in the final analysis. The data were analyzed using *R*. Descriptive statistics were used to summarize the sample's demographic characteristics, reporting means, and standard deviations. Additionally, a series of OLS and OLR models with robust standard errors were performed to assess the influence of participant demographics and gender bias on candidate assessment scores. The survey experiment was structured to isolate the effects of candidate gender and qualifications while capturing nuanced differences in participant evaluations. Combined with demographic data, this approach allows for a multifaceted analysis of potential biases and perceived discrimination in STEM hiring contexts. The experimental procedures are described in Subsection 3.1.1, while the structure of the survey instrument and question formats is outlined in Section 3.2.

3.1.1 Experimental Setup and Procedure

Building on the structure outlined in the Research Design section, the experiment was designed to examine how

[2] S&P Global, *Women in Leadership: What's the Holdup?*, accessed March 22, 2025, https://www.spglobal.com/esg/insights/featured/special-editorial/women-in-leadership-what-s-the-holdup?utm_source=chatgpt.com.

participants evaluate job candidates and reflect on gender-related challenges in hiring. Each participant took on the role of a hiring manager and reviewed a resume on various aspects related to their suitability for the job. To keep the evaluation process realistic, participants were told that the resumes belonged to real applicants and that their feedback would help improve hiring decisions. After completing the evaluation, participants answered a post-experiment questionnaire focused on perceptions of workplace discrimination. The survey was conducted online via *Qualtrics*. The following section explains how the resumes were developed to systematically vary candidate gender and qualification strength. The template for the resumes was selected based on online research as one of the most popular templates in the job market. This approach replicates the methodology used by Bertrand and Mullainathan (2004), who constructed resumes based on real templates and systematically varied candidate qualifications to analyze hiring discrimination. Like their study, resume differentiation in this experiment was achieved through variations in academic performance, technical expertise, and certification levels.

The candidate profiles were presented in a standardized format, including consistent sections on education, skills, work experience, and certifications. The content emphasized a STEM background with a focus on electrical engineering, complemented by some coursework in business-related subjects. Each profile reflected a balance of theoretical knowledge and practical experience, typical of entry-level applicants in a technical-oriented field. To show differences in candidate qualifications, their GPAs were adjusted to reflect different levels of academic and analytical ability. Candidates with lower GPA scores were portrayed as having weaker theoretical skills, while those with higher GPAs emphasized stronger academic performance. *MATLAB*, an important tool for electrical engineers, was added to the strong resumes to highlight stronger technical expertise (Stanculescu et al., 2009). Weak resumes included fewer technical skills. Differences in certifications (e.g. IELTS and GMAT) and scores were designed to indicate varying levels of ambition and competence. Strong resumes included higher scores and additional certifications. The experience levels of the candidates were differentiated by length of employment and range of responsibilities. Strong profiles showed longer work experience and more hands-on tasks, suggesting a higher level of professional development. Since applicants with more work experience are often seen as more competent, experience was used as a key difference between profiles (Paloniemi, 2006). The described candidate resumes used in the experiment are provided in Appendix B.

The fictitious candidates were named Sophie and Max Schneider. These names were carefully chosen to convey a clear gender identity without causing confusion, which was crucial, given the study's focus on gender perceptions in hiring decisions. Another deciding factor in selecting these names was their popularity among German newborns between 1997 and 2005, which aligned with the expected age range of recent university graduates.[3] This ensured that the names would feel familiar and realistic to survey participants. Schneider, one of the most common surnames in Germany, particularly in the southern regions, was selected to further enhance the authenticity of the candidate profiles.[4]

Participants received an introduction explaining the purpose of the study – to simulate a hiring decision – without revealing its specific focus on gender bias to avoid influencing answers. They were instructed to carefully review the candidate's resume and assess them on key dimensions, including hireability, competence, and potential starting salary. After evaluating the candidate, participants completed a short follow-up questionnaire on gender discrimination. On average, participants spent approximately 8 to 10 minutes completing the entire survey, including both the resume evaluation and post-experiment questionnaire (Kost & da Rosa, 2018). This duration is in line with the best practices for online surveys, where engagement typically declines after 10 minutes.[5] The experiment was carefully controlled to ensure the data could be used to draw valid conclusions about participants' views on gender-related challenges in hiring.

3.2 Survey Design

The survey is designed to collect data in a structured format, comprising three main sections: Demographic Questions, Experimental Evaluation, and Perception of Discrimination Questions. Each section is designed to ensure comprehensive data collection while maintaining data quality and participant engagement. The full version of the survey is provided in Appendix C.

In the first section, demographic data is collected to understand the background and characteristics of the participants. These include 7 questions based on age, gender, country of origin, pursued university degree, field of study, previous field of study, and work experience. Age was included to analyze possible generational differences, such as attitudes toward gender roles (Meagher & Shu, 2019; Yigit & Ali, 2024). Country of origin was included to capture cultural and societal influences, recognizing the diversity of norms across countries, which might be an interesting topic for future research (Cuddy et al., 2015). Questions on current and previous fields of study helped compare STEM and non-STEM participants, focusing on discipline-specific trends. Work experience provided insight into the extent to which participants were affected by workplace dynamics

[3] The names Max and Sophie were among the top 20 most popular baby names in Germany during the 2000s, accessed March 19, 2025, https://www.babymed.com/unassigned/top-german-baby-names-germany-2000s.

[4] Wikipedia, *List of the Most Common Surnames in Germany*, last modified March 19, 2025, https://en.wikipedia.org/wiki/List_of_the_most_common_surnames_in_Germany.

[5] SurveyMonkey, *Survey Completion Times: How Long Do People Take to Complete a Survey?*, accessed March 19, 2025, https://www.surveymonkey.com/curiosity/survey_completion_times/.

that may have influenced their perceptions of bias (Funk & Parker, 2018).

Demographic questions helped to group independent variables and conduct descriptive statistics to better characterize the sample (e.g., gender distribution, average age). In addition, these variables were used to analyze group differences and find relationships between independent and dependent variables, such as the possible influence of gender on hireability ratings. Special attention was given to gender, work experience, and fields of study, as these were the main interests in identifying trends and understanding how background factors contribute to participants' perceptions and evaluations. At this stage, a factual attention check question is included to ensure data quality by identifying inattentive or low-effort responses that could compromise the reliability and integrity of the results (Kung et al., 2018; Muszyński, 2023).

In the second section, participants are introduced to the experimental part of the survey. This study used an experimental audit research design typically used in field studies to analyze discrimination, which simulated hiring scenarios to examine gender bias in candidate evaluations of university students. Students were asked to rate applicants for entry-level electrical engineering roles. The electrical engineering role was chosen with the goal that resumes did not contain specific qualifications unfamiliar to students, as most of the respondents had STEM backgrounds. Students were provided with clear instructions on how to evaluate a candidate's CV, emphasizing that they should carefully review the qualifications and assess them based on several criteria. To improve the quality of responses, participants were asked to give honest evaluations, regardless of whether they thought the application was strong or weak. A request for commitment was also included, as research has shown that securing commitment from participants promotes greater engagement and increases the reliability of the data.[6]

The Candidate Evaluation section assessed whether students exhibit gender biases when evaluating job applicants based on resumes. Specifically, participants rated candidates' competence, hireability, salary expectations, and willingness to assist them in a professional or academic setting. This evaluation provided insights into whether gender influences students' hiring-related perceptions and decision-making. The assessment questions were adapted from Moss-Racusin et al. (2012), who investigated gender bias in hiring decisions among science faculty. This study included the competence evaluation criteria, which is a key determinant of hiring decisions and is often influenced by implicit gender biases (Heilman, 2012). Participants rated the applicant's competence using three questions:

1. Did the applicant strike you as competent?
2. How likely is it that the applicant has the necessary skills for this job?
3. How qualified do you think the applicant is?

Responses were recorded on a 5-point Likert scale (1 = Not at all, 5 = Very much), with higher scores indicating greater perceived competence.

Hireability was evaluated to determine whether the gender of an applicant influenced participants' likelihood of selecting them for a job. Prior research has shown that gender bias often manifests in lower hireability ratings for equally qualified female candidates (Rudman & Glick, 2001). To assess this, participants responded to three Likert-scale questions:

1. How likely would you be to hire the applicant?
2. How likely would you be to recommend the applicant for the job?
3. How likely is it that the applicant would actually be hired for the position?

These items were rated on a 5-point Likert scale (1 = Not at all likely, 5 = Very likely), with higher scores reflecting greater perceived hireability.

Salary allocation indicates how much the evaluator values the candidate's contributions (Gerhart & Rynes, 2003). To examine this, participants were asked: "*If you had to choose a starting salary for this applicant, which of the following would you select?*". Response options ranged from less than € 35,000 to more than € 70,000 per year (before taxes). This salary range was chosen based on market research, which identified € 50,000 as the average entry-level salary for an electrical engineering position.[7] This measure helped determine whether participants were assigned different salaries based on the applicant's gender, reflecting perceived value and career potential.

Professional support was also adapted from the mentoring measure used by Moss-Racusin et al. (2012). Unlike faculty hiring managers, students are not in a position to mentor applicants. Therefore, the mentoring question was changed to better reflect student evaluators' experiences. Instead of asking whether participants would mentor the applicant, the question was: "*If you were working alongside this candidate at your university or in a similar professional setting, how likely would you be to offer assistance if they were struggling with a challenging concept or project?*". The response was recorded using a 5-point Likert scale ranging from Extremely unlikely to Extremely likely.

In the third section, participants answer questions related to their perception of gender-specific challenges in hiring, mentoring, and promotion. More specifically, they answered questions related to perceived discrimination against

[6] Qualtrics, *Attention Checks and Data Quality: How to Improve Survey Responses*, accessed March 19, 2025, https://www.qualtrics.com/blog/attention-checks-and-data-quality/.

[7] PayScale, *Electrical Engineer Salary in Germany*, accessed March 19, 2025, https://www.payscale.com/research/DE/Job=Electrical_Engineer/Salary.

women, discrimination against men, and discrimination that might affect them personally (self-related discrimination). These 5-point Likert-scale questions from Strongly Disagree to Strongly Agree are designed to measure participants' beliefs and attitudes toward gender discrimination and bias within workplace settings, particularly in STEM fields. These questions were adapted from Sipe et al. (2009), who previously examined university students' perceptions of gender discrimination in the workplace. These questions were modified to focus specifically on STEM fields, aligning with the central research question of this study on gender bias in STEM hiring. For example, the original item "Women/Men are less likely to be promoted due to their gender" was adapted to "Women/Men are less likely to be promoted within STEM fields due to their gender". Such modifications ensured that the questions directly reflected the hiring environment relevant to the participants and the study's objectives. The specific questions included:

1. Women are less likely to be hired within STEM fields due to their gender / Men are less likely to be hired within STEM fields due to their gender.

2. Women are less likely to be promoted within STEM fields due to their gender/ Men are less likely to be promoted within STEM fields due to their gender.

3. Women face more challenges in networking and mentorship opportunities within STEM fields due to their gender / Men face more challenges in networking and mentorship opportunities within STEM fields due to their gender.

4. Women are more likely to be underpaid within STEM fields compared to men with similar qualifications / Men are more likely to be underpaid within STEM fields compared to women with similar qualifications.

3.3 Data Collection

Data were collected over a two-month period, from 2 December 2024 to 3 February 2025, using an online survey hosted on the *Qualtrics* platform. Participants accessed the survey through an automatically generated anonymous link, which was distributed via multiple channels, including the university mailing list, *Moodle* platform, class announcements, and *WhatsApp* groups. Additionally, the university's Student Counseling department assisted with promoting the survey to students.

Participants were invited to take part in the study through a message that explained the survey was anonymous and confidential (see Appendix B). After clicking the link, they saw a short description of the study, which included its purpose, the voluntary nature of participation, and the estimated time to complete it. By continuing with the survey, participants gave their consent to take part and allowed their responses to be used for research. *Qualtrics* settings ensured that no personal or sensitive information was collected. Participants could leave the survey at any time without any consequences. To eliminate bias in resume distribution, a randomization tool in *Qualtrics* was used to ensure that each participant received only one of the four possible resumes.

A total of 127 responses were recorded during the data collection period. Of these, 92 were fully completed and included in the final analysis. Participants who did not answer all survey questions, including one who skipped the final item, were excluded. To assess the comparability of the groups, a balancing table was constructed. Table 1 presents descriptive statistics across the four experimental conditions: strong/weak resumes and female/male candidate gender. It includes demographic characteristics of the participants (age, gender, work experience, and STEM background), as well as information on education and region of origin. The results indicate that the groups are generally well balanced, with only minor differences that do not suggest any systematic bias in participant allocation. This supports the conclusion that the random assignment procedure was successful and allows for valid data analysis to follow.

4 Data Analysis

This chapter outlines the analytical strategy used to examine participants' evaluations of hypothetical job candidates and their perceptions of gender-based discrimination in hiring. The analysis includes both descriptive and inferential components, beginning with an overview of variable construction, followed by descriptive insights into perceived discrimination, and concluding with a series of regression models designed to test the main question of the study.

The analysis focuses on four primary outcome variables: competence, hireability, professional support, and recommended starting salary. Two key experimental variables, the candidate's gender, and qualification, form the core of the independent variables. Candidate gender is coded as a binary variable (0 = female, 1 = male), with female as the baseline category. Candidate qualification is likewise binary, where strong candidates are coded as 1 and weak candidates as 0 (reference category). Several control variables are included to account for participants' demographic characteristics. These include gender (binary variable), age (categorized with age groups), country of origin (grouped into broader regional categories), academic background (STEM vs. non-STEM), and work experience (with grouped categories for years of experience). All categorical variables are dummy-coded for regression analyses to isolate the effects of candidate characteristics while accounting for participant background.

To complement the analysis of candidate evaluations, the study also explores how students perceive gender-based discrimination in hiring and workplace dynamics. These attitudes provide a critical context for interpreting evaluation behaviors. Three perception variables were used for this purpose: Discrimination Against Self, Discrimination Against Men and Discrimination Against Women. These

Table 1: Balancing Table of Resume Types

Variable	Strong Female	Strong Male	Weak Female	Weak Male
Mean Age	23.48	24.27	27.13	25.55
	(3.96)	(3.60)	(5.36)	(4.53)
Mean Gender (1 = Male)	0.43	0.43	0.65	0.48
	(0.51)	(0.51)	(0.49)	(0.51)
Mean Work Experience (1 = Yes)	0.83	0.87	0.86	0.95
	(0.39)	(0.34)	(0.35)	(0.21)
Bachelor's Degree (%)	65.22%	52.17%	26.09%	43.48%
Master's Degree (%)	34.78%	47.83%	69.57%	43.48%
PhD (%)	0.00%	0.00%	4.35%	13.04%
Mean STEM (1 = STEM)	0.52	0.52	0.65	0.61
	(0.51)	(0.51)	(0.49)	(0.50)
Regional Distribution				
Caucasus (%)	43.48%	65.22%	21.74%	13.04%
Western Europe (%)	4.35%	8.70%	21.74%	17.39%
Eastern Europe (%)	13.04%	8.70%	17.39%	0.00%
MENA (%)	17.39%	13.04%	21.74%	52.17%
Asia (%)	13.04%	4.35%	17.39%	8.70%
Americas (%)	8.70%	0.00%	0.00%	8.70%
Sample Size	23	23	23	23

Notes: This table presents the balancing of the four resume types used in the experiment. Variables include key participant characteristics such as age, gender, work experience, education level, and STEM, along with regional distribution. STEM is coded as 1 if the participant is from a STEM field and 0 otherwise. MENA refers to the Middle East and North Africa region. Sample sizes represent the number of respondents per resume type.

questions asked about perceived bias in hiring, promotion, salary offers, and mentoring or networking opportunities. While these variables are not included in the regression models as outcomes, they offer descriptive insights into how gender bias is perceived by different student subgroups. Group-based comparisons (e.g., by gender, study field, or work experience) help identify patterns in perception that may inform or parallel evaluation behaviors.

To investigate how candidate characteristics and participant demographics influence candidate evaluations, a series of regression models were estimated. Competence, hireability, and professional support were analyzed using OLS regression. These outcomes were measured on 5-point Likert scales and treated as continuous variables to simplify analysis and increase statistical power. However, for salary recommendations, an OLR model was used. This choice reflects the ordered nature of the salary data, which was collected in categories (e.g., € 35k–€ 45k, € 45k–€ 55k). OLR preserves the ranking of these categories without assuming equal distances between them, making it more suitable than OLS in this case. Robust standard errors were used in all models to correct for potential heteroskedasticity and ensure reliable statistical inference. This approach allows for appropriately analyzing both interval-like and ordinal outcomes within the survey experiment.

The central regression model is specified as follows:

$$\begin{aligned} Y_i = \beta_0 + \beta_1 \cdot \text{CandGender} + \beta_2 \cdot \text{CandQual} \\ + \beta_3 \cdot (\text{CandGender} \times \text{CandQual}) + X\gamma + \epsilon_i \end{aligned} \tag{1}$$

In this specification, the outcome variable Y_i represents one of four participant evaluations of a hypothetical job candidate: competence, hireability, professional support, or recommended starting salary. β_0 captures the baseline evaluation score for a female candidate with weak qualifications, assessed by a participant from the reference group. CandGender is a binary indicator equal to 1 if the candidate is male and 0 if female. CandQual equals 0 if the CV is weak and 1 if strong. The interaction terms CandGender × CandQual are included to assess whether the effects of candidate gender vary depending on the qualification strength of the respondent. X is a vector of control variables capturing participant characteristics (gender, age category, country of origin, study field, and work experience), γ is the associated vector of coefficients, and ϵ_i is the error term. The results of analyses are presented and interpreted in Chapter 5.

5 Results

This chapter presents the findings from both the descriptive analysis and the regression models introduced in Chapter 4. It begins by exploring how participants perceive gender-based discrimination in STEM fields. The second part of the chapter examines how candidate characteristics and participant demographics influence participants' decision-making in the simulated hiring scenarios.

5.1 Descriptive Statistics

This section examines students' perceptions of gender-based discrimination across three dimensions: discrimination against women, against men, and against oneself. Results vary by gender, work experience, academic background, and regional origin. Full descriptive statistics are provided in Table A 6. Discrimination against women was perceived as the most prevalent dimension of all. Furthermore, female students reported substantially higher awareness in this regard than male students (M = 3.54 vs. 2.65), and non-STEM participants also showed stronger agreement. These results align with prior research suggesting that women may be more attuned to structural gender inequalities. In terms of regional differences, students from the Americas expressed the highest levels of perception with respect to bias against women, while those from Western Europe were the least aware of the matter. In contrast, discrimination against men received the lowest overall ratings, though male students rated it slightly higher than female students (M = 2.26 vs. 2.07). Perceptions were somewhat elevated among STEM students and those with 5–7 years of work experience. When it comes to discrimination perceived on a personal level, students with 1–3 years of work experience reported the highest scores. Female students and those from non-STEM fields again reported greater concern, while participants with more than seven years of experience expressed the lowest levels. These results suggest that early-career students may feel more vulnerable to bias, whereas longer experience may be associated with greater confidence or adaptation to workplace dynamics. Taken together, these findings highlight the diversity of students' beliefs and expectations surrounding gender discrimination in STEM. The differences observed across gender, experience, academic field, and regional background provide important context for the experimental results presented in the next section, where students' actual evaluation behaviors are examined.

5.2 Regression Analyses

This section presents the results of regression models assessing how candidate gender and qualifications, along with participant demographics, influenced evaluations across four outcomes: competence, hireability, professional support, and salary recommendations. Results are grouped by a dependent variable, the key coefficients of each model are discussed in the text, and detailed regression tables are provided below. The regression analyses were performed in multiple stages. First, bivariate models estimated the effect of candidate gender or qualifications on the outcome variable without including participant controls. These were followed by multivariate models that controlled for respondent characteristics such as age, gender, origin, academic background, and work experience. Finally, the full models included interaction terms between candidate gender and qualifications to assess whether evaluation patterns change depending on the strength of the candidate's profile.

Candidate Competence

As shown in Table 2, none of the experimental manipulations significantly affected competence scores. In the bivariate model, male candidates were rated similarly to female candidates ($\beta = -0.13$), and this remained the case after including participant covariates ($\beta = -0.11$). Candidate qualifications also had no significant influence on perceived competence either in isolation or when adjusting for covariates. The interaction between gender and qualifications was consistently non-significant, including in the full model. In contrast, some participant characteristics did shape competence evaluations. Participants aged 23–27 and over 33 provided significantly higher ratings than those under 23. Additionally, respondents from Eastern Europe gave lower ratings to all candidates, suggesting regional variation in evaluation standards. No other demographic covariates had a notable impact.

Candidate Hireability

Table 3 illustrates regression estimates for hireability evaluations. Candidate gender significantly influenced perceived hireability in the bivariate model, with male candidates rated less hireable than female candidates ($\beta = -0.326$). After controlling for participant demographics, this effect remained marginally significant ($\beta = -0.276$), indicating a potential bias favoring female candidates. Candidate qualifications did not significantly predict hireability in any model. However, the interaction between candidate gender and qualification in Model 5 resulted in a marginally significant positive coefficient ($\beta = 0.565$), suggesting strong male candidates may be evaluated more favorably than weak male candidates. This effect weakened in the full model ($\beta = 0.404$) and was no longer significant. Among participant covariates, those aged 23–27 and those over 33 gave higher hireability scores. Additionally, participants from Eastern Europe tended to assign lower ratings. Other demographic variables were not statistically significant.

Professional Support

Table 4 presents the results of OLS regression models of participants' willingness to offer professional support to candidates. In Model 1, which includes only candidate gender, the coefficient for male candidates is negative ($\beta = -0.109$) but not statistically significant, indicating that gender alone does not predict the level of support. When demographic covariates are added in Model 2, the coefficient remains negative ($\beta = -0.147$) and non-significant. In Model 3, the effect of candidate qualifications alone on professional support did not show a significant correlation. However, when covariates are added in Model 4, age shows a significant predictor, especially for older participants. Model 5 introduces the interaction between candidate gender and qualifications. The main effect of being a strong candidate is marginally positive ($\beta = 0.435$), but this is moderated by a marginally negative interaction term ($\beta = -0.652$), suggesting that while strong candidates generally receive more support, this effect may not apply equally across genders. In the full model (Model 6), the interaction term remains negative ($\beta = -0.728$) and marginally significant, indicating that strong male candidates receive less support than expected based on the additive effects of gender and qualifications.

Table 2: OLS Regression – Dependent Variable: Competence Score

Variable	(1)	(2)	(3)	(4)	(5)	(6)
Candidate Male	-0.130	-0.112	–	–	-0.174	-0.064
	(0.165)	(0.153)			(0.256)	(0.207)
Candidate Strong	–	–	0.087	-0.0004	0.043	0.044
			(0.165)	(0.173)	(0.230)	(0.216)
Male × Strong	–	–	–	–	0.087	-0.093
					(0.333)	(0.274)
Constant	3.630***	3.759***	3.522***	3.454***	3.609***	3.498***
	(0.114)	(0.169)	(0.127)	(0.204)	(0.151)	(0.210)
Covariates (Z)	–	✓	–	✓	–	✓
R-squared	0.007	0.343	0.003	0.339	0.011	0.344
Observations	92	92	92	92	92	92

Notes: Robust standard errors in parentheses. *p<0.10, **p<0.05, ***p<0.01.

Table 3: OLS Regression – Dependent Variable: Hireability Score

Variable	(1)	(2)	(3)	(4)	(5)	(6)
Candidate Male	-0.326*	-0.276*	–	–	-0.609*	-0.487*
	(0.160)	(0.163)			(0.256)	(0.236)
Candidate Strong	–	–	0.109	0.038	-0.174	-0.145
			(0.164)	(0.189)	(0.196)	(0.210)
Male × Strong	–	–	–	–	0.565*	0.404
					(0.318)	(0.284)
Constant	3.783***	3.825***	3.565***	3.634***	3.870***	3.880***
	(0.098)	(0.215)	(0.134)	(0.280)	(0.130)	(0.273)
Covariates (Z)	–	✓	–	✓	–	✓
R-squared	0.044	0.250	0.005	0.223	0.082	0.265
Observations	92	92	92	92	92	92

Notes: Robust standard errors in parentheses. *p<0.10, **p<0.05, ***p<0.01.

Salary Recommendations

Table 5 presents the results of OLR regression models analyzing salary expectations for the presented candidates. In models 1 and 3, neither candidate gender nor candidate qualification has a statistically significant effect on the recommended salary category. However, the coefficient for a strong candidate is positive ($\beta = 0.530$) in Model 3, indicating a tendency towards higher salary suggestions, though this effect is not statistically significant. When controlling for demographic covariates, strong candidates are significantly more likely to receive higher salary recommendations ($\beta = 0.849$). This effect persists in the full model, although it is no longer statistically significant. In Model 5, we observe a positive, but non-significant interaction between candidate gender and qualifications. Later, the interaction becomes marginally significant in the full model, suggesting that male candidates with strong qualifications may receive higher salary recommendations than their female counterparts. These findings suggest that while gender and qualifications alone do not strongly predict salary recommendations, their interaction and the characteristics of the evaluators themselves may play an influential role. An overview of these findings is depicted in Figures B6–B8.

Influence of Participant Demographics

To explore whether certain participant characteristics influenced how candidates were evaluated, I ran an additional set of models using only demographic variables. The results are presented in Tables A 7 and A 8 in the Appendix A. Participants aged 23–27 ($\beta = 0.557$) and especially those over 33 ($\beta = 0.916$) gave significantly higher competence ratings compared to the youngest group. The same age groups also showed more willingness to offer professional support, with the strongest effect observed among those over 33 ($\beta = 1.605$). Age also had a positive effect on

Table 4: OLS Regression – Dependent Variable: Professional Support Score

Variable	(1)	(2)	(3)	(4)	(5)	(6)
Candidate Male	-0.109	-0.147	–	–	0.217	0.231
	(0.202)	(0.227)			(0.266)	(0.329)
Candidate Strong	–	–	0.109	0.080	0.435*	0.417
			(0.202)	(0.235)	(0.273)	(0.295)
Male × Strong	–	–	–	–	-0.652*	-0.728
					(0.402)	(0.452)
Constant	4.000***	3.637***	3.891***	3.486***	3.783***	3.411***
	(0.139)	(0.307)	(0.133)	(0.332)	(0.198)	(0.368)
Covariates (Z)	–	✓	–	✓	–	✓
R-squared	0.003	0.214	0.003	0.210	0.035	0.245
Observations	92	92	92	92	92	92

Notes: Robust standard errors in parentheses. $^{*}p<0.10$, $^{**}p<0.05$, $^{***}p<0.01$.

Table 5: Ordinal Logistic Regression – Salary Recommendations

Variable	(1)	(2)	(3)	(4)	(5)	(6)
Candidate Male	-0.019	0.216	–	–	-0.453	-0.472
	(0.378)	(0.414)			(0.531)	(0.607)
Candidate Strong	–	–	0.531	0.849*	0.093	0.224
			(0.383)	(0.468)	(0.534)	(0.605)
Male × Strong	–	–	–	–	0.892	1.435*
					(0.762)	(0.866)
Covariates (Z)	–	✓	–	✓	–	✓
AIC	259.96	266.55	258.03	263.48	260.65	264.31
Observations	92	92	92	92	92	92

Notes: Standard errors in parentheses. $^{*}p<0.10$, $^{**}p<0.05$, $^{***}p<0.01$. *AIC = Akaike Information Criterion; lower values indicate better model fit.*

hireability ratings, although the result for participants over 33 was only marginally significant (β = 0.924). Regional origin also played a role in shaping evaluation behavior. Respondents from Eastern Europe consistently rated candidates lower in terms of competence (β = -0.732) and hireability (β=-0.454). A negative trend was also observed among participants from Asia, particularly in competence and support ratings. On the other hand, students from Asia and the Americas were more likely to assign higher salary categories, indicating potential regional differences in perceived salary expectations. Work experience was especially relevant for salary evaluations. Those with five to seven years (β = -2.941) and more than seven years (β = -2.318) of experience tended to suggest lower starting salaries, which may reflect more cautious or realistic expectations. Other demographic variables, including participant gender and academic background, did not show a consistent or statistically significant influence on candidate evaluations.

5.3 Robustness Check

To make sure that the main results are reliable, I repeated the full model (Model 6) using only participants with a STEM academic background (n = 53). This group is particularly relevant because the study focuses on perceptions of hiring in STEM fields. The goal was to test whether the patterns found in the overall sample also appear among those most likely to enter STEM careers. The results are summarized in Table A 9.

Across all outcome variables, the main effects of candidate gender, qualification, and their interaction were not statistically significant. For competence ratings, however, demographic characteristics continued to play a role: older participants (33+) gave higher competence scores, while participants from Eastern Europe and Asia gave lower scores. A similar pattern was observed in hireability ratings, where no significant treatment effects emerged, although participants from Eastern Europe tended to assign slightly lower scores. For professional support, the only notable result was that older participants were more willing to offer help to struggling candidates. Finally, salary recommendations showed a marginally significant interaction effect ($\beta = 2.55$), indicating a possible tendency to favor highly qualified male candidates. Age and origin again played a role, with older participants and those from Eastern Europe, Asia, and the Americas suggesting higher starting salaries.

Overall, the robustness check using the STEM-only sample confirms that the direction of effects remains similar to the full sample. While most treatment effects are not statistically significant in this subgroup, the influence of participant demographics, particularly age and region of origin, remains evident. These findings support the stability of the results and suggest that background characteristics continue to shape perceptions, even within a more homogeneous STEM population.

6 Discussion and Conclusion

To investigate how university students perceive gender in hiring contexts, this study used an experimental design with manipulated CVs to explore how they evaluate applicants based on standard hiring criteria. The analysis considered both the direct effects of candidate gender and interaction effects with qualification strength, as well as the influence of respondent demographic characteristics. Despite expectations based on generally accepted findings in labor economics and psychology, the study did not reveal a consistent preference for male candidates based on assessment results. In simpler models, male candidates were rated lower on hireability.

The presented results raise an interesting question about how this shift may have occurred among students. One plausible explanation is that the academic and cultural environment of modern universities, where diversity and inclusion are actively promoted, especially in STEM fields, is beginning to create a more egalitarian mentality[8] in students. Perhaps the impact of inclusive discourse has influenced their values, making them less likely to apply traditional gender stereotypes in their evaluations. Social desirability bias may still have influenced responses, as students could have unconsciously adjusted their answers to match internalized social norms about fairness and equality, even without knowing the study's true focus (Fisher, 1993; Krumpal, 2013). While the direct effects of candidate gender were limited, interaction terms revealed more nuanced dynamics. Strong male candidates were marginally more likely to receive higher salary recommendations but less likely to receive professional support compared to other groups. This may reflect typical assumptions in male-dominated fields: strong men are expected to be competent and independent and, therefore, not need help. This pattern aligns with the shifting standards framework (Biernat & Kobrynowicz, 1997), which suggests that people do not use a single, fixed standard when evaluating others. Instead, they adjust their expectations based on social group stereotypes, for example, assuming women in STEM are less common, they might be judged as "exceptional" for meeting the same criteria expected of men. As a result, men might receive less support, not because of bias but because of underlying expectations that they should manage on their own in these roles.

Another key insight is that the evaluators' background, especially their age and region of origin, consistently influenced how they rated candidates across all outcomes. Regional differences, especially among students from Eastern Europe and Asia, may reflect cultural norms around hierarchy, competition, or gender roles. For example, stricter academic standards in some regions may influence how competence is rated, while conservative gender norms might shape attitudes toward female candidates. Moreover, students with more work experience recommended lower starting salaries, suggesting a more labor-market-aware perspective. This indicates that participants' views on fairness and value may have been shaped not only by bias but also by their practical experience and knowledge of how the job market works. In addition to behavioral measures, the study explored students' perceptions of gender-based discrimination. Perceptions of self-related discrimination were strongest among early-career students, suggesting that perceived vulnerability to bias may decline with experience.

While the study contributes valuable insights, several limitations must be noted. First, the sample size (n = 92) limits the ability to detect minor effects and fully explore subgroup variation. Second, as in most online experiments, participants faced no real consequences for their decisions. This may reduce ecological validity[9], as real-life hiring involves risks, incentives, and institutional norms that shape behavior in ways experiments cannot fully simulate. Third, although the average survey completion time was within an acceptable range (8–10 minutes), some participants completed it in under 7 minutes, raising concerns about attentiveness and the depth of engagement with the evaluation task. Fourth, while the resumes were designed to reflect a neutral, entry-level electrical engineering position, participants from non-engineering backgrounds (e.g., business and social sciences) may have lacked the context to accurately interpret technical qualifications despite efforts to make the content broadly understandable. Finally, the study focused solely on gender and qualification strength.

These limitations open several promising directions for future research. One important avenue is to explore intersectional dimensions – such as ethnicity, age of the applicant, or motherhood/fatherhood signals – to better understand how multiple identities may compound bias in hiring evaluations. Additionally, longitudinal research could examine how students' attitudes evolve as they transition into professional roles, revealing whether the relatively egalitarian

[8] According to the Oxford English Dictionary, egalitarianism is "the belief in or advocacy of the principle that all people are equal and deserve equal rights and opportunities."

[9] Ecological validity refers to how well the results of a study can be applied to real-life settings and behaviors outside the experimental environment.

patterns observed in academic settings persist or shift under real-world workplace pressures. Cross-cultural studies could also provide valuable insights by comparing results across different countries or academic systems, helping to identify which institutional features most influence bias-related behavior. Finally, future work should investigate the potential of educational interventions to reduce gender bias and promote more equitable evaluation practices. Can targeted training in bias awareness meaningfully change how people evaluate others? And are such effects durable over time, particularly once individuals enter professional roles and face more complex decision-making environments?

Taken together, the findings suggest that, while students are not immune to bias, their judgments reflect a shift toward greater awareness and caution. Compared to previous generations, today's students seem to be more sensitive to issues of equity and inclusion. However, awareness alone may not eliminate deeper social expectations or implicit assumptions. As bias appears increasingly context-sensitive and less predictable, efforts to combat it must also evolve through education, reflection, and policies that take into account the complex ways in which bias operates in practice.

References

Alan, S., Ertac, S., Kubilay, E., & Loranth, G. (2019). Understanding Gender Differences in Leadership. *The Economic Journal*, *130*(626), 263–289. https://doi.org/10.1093/ej/uez050

Babcock, L., Laschever, S., Gelfand, M., & Small, D. (2003). Nice Girls Don't Ask. *Harvard Business Review*, *81*(10), 14–14.

Bertrand, M., & Mullainathan, S. (2004). Are Emily and Greg More Employable Than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination. *American Economic Review*, *94*(4), 991–1013.

Biernat, M., & Kobrynowicz, D. (1997). Gender- and Race-Based Standards of Competence: Lower Minimum Standards but Higher Ability Standards for Devalued Groups. *Journal of Personality and Social Psychology*, *72*(3), 544.

Correll, S. J. (2001). Gender and the Career Choice Process: The Role of Biased Self-Assessments. *American Journal of Sociology*, *106*(6), 1691–1730. Retrieved March 21, 2025, from http://www.jstor.org/stable/10.1086/321299

Cuddy, A. J., Wolf, E. B., Glick, P., Crotty, S., Chong, J., & Norton, M. I. (2015). Men as Cultural Ideals: Cultural Values Moderate Gender Stereotype Content. *Journal of Personality and Social Psychology*, *109*(4), 622.

Evagorou, M., Puig, B., Bayram, D., Janeckova, H., & European Commission and Directorate-General for Education, Youth, Sport and Culture. (2024). *Addressing the Gender Gap in STEM Education across Educational Levels – Analytical Report*. Publications Office of the European Union. https://doi.org/10.2766/260477

Fisher, R. J. (1993). Social Desirability Bias and the Validity of Indirect Questioning. *Journal of Consumer Research*, *20*(2), 303–315.

Friedmann, E., & Efrat-Treister, D. (2023). Gender Bias in STEM Hiring: Implicit In-Group Gender Favoritism Among Men Managers. *Gender & Society*, *37*(1), 32–64. https://doi.org/10.1177/08912432221137910

Funk, C., & Parker, K. (2018). Women and Men in STEM Often at Odds over Workplace Equity [https://www.pewresearch.org/social-trends/2018/01/09/women-and-men-in-stem/].

Gerhart, B., & Rynes, S. (2003). *Compensation: Theory, Evidence, and Strategic Implications*. Sage.

Goldin, C., & Rouse, C. (2000). Orchestrating Impartiality: The Impact of "Blind" Auditions on Female Musicians. *American Economic Review*, *90*(4), 715–741.

Hardy III, J. H., Tey, K. S., Cyrus-Lai, W., Martell, R. F., Olstad, A., & Uhlmann, E. L. (2022). Bias in Context: Small Biases in Hiring Evaluations Have Big Consequences. *Journal of Management*, *48*(3), 657–692. http://www.jstor.org/stable/10.1086/321299

Heilman, M. E. (2012). Gender Stereotypes and Workplace Bias. *Research in Organizational Behavior*, *32*, 113–135.

Holman, L., Stuart-Fox, D., & Hauser, C. E. (2018). The Gender Gap in Science: How Long until Women Are Equally Represented? *PLoS Biology*, *16*(4), e2004956.

Kost, R. G., & da Rosa, J. C. (2018). Impact of Survey Length and Compensation on Validity, Reliability, and Sample Characteristics for Ultrashort-, Short-, and Long-Research Participant Perception Surveys. *Journal of Clinical and Translational Science*, *2*(1), 31–37. https://doi.org/10.1017/cts.2018.18

Krumpal, I. (2013). Determinants of Social Desirability Bias in Sensitive Surveys: A Literature Review. *Quality & Quantity*, *47*(4), 2025–2047.

Kung, F. Y., Kwok, N., & Brown, D. J. (2018). Are Attention Check Questions a Threat to Scale Validity? *Applied Psychology*, *67*(2), 264–283. https://doi.org/10.1111/apps.12108

Meagher, K. D., & Shu, X. (2019). Trends in US Gender Attitudes, 1977 to 2018: Gender and Educational Disparities. *Socius*, *5*, 2378023119851692.

Milkman, K. L., Akinola, M., & Chugh, D. (2015). What Happens Before? A Field Experiment Exploring How Pay and Representation Differentially Shape Bias on the Pathway into Organizations. *Journal of Applied Psychology*, *100*(6), 1678.

Moss-Racusin, C. A., Dovidio, J. F., Brescoll, V. L., Graham, M. J., & Handelsman, J. (2012). Science Faculty's Subtle Gender Biases Favor Male Students. *Proceedings of the National Academy of Sciences*, *109*(41), 16474–16479.

Muszyński, M. (2023). Attention Checks and How to Use Them: Review and Practical Recommendations. *Research and Methods*, *32*(1), 3–38. https://doi.org/10.18061/ask.v32i1.0001

Neumark, D. (2018). Experimental Research on Labor Market Discrimination. *Journal of Economic Literature*, *56*(3), 799–866.

Paloniemi, S. (2006). Experience, Competence and Workplace Learning. *Journal of Workplace Learning*, *18*(7/8), 439–450. https://doi.org/10.1108/13665620610693006

Quadlin, N. (2018). The Mark of a Woman's Record: Gender and Academic Performance in Hiring. *American Sociological Review*, *83*(2), 331–360.

Rich, J. (2014). *What Do Field Experiments of Discrimination in Markets Tell Us? A Meta Analysis of Studies Conducted since 2000* (tech. rep.). IZA Discussion Papers.

Rudman, L. A., & Glick, P. (2001). Prescriptive Gender Stereotypes and Backlash toward Agentic Women. *Journal of Social Issues*, *57*(4), 743–762. https://doi.org/10.1111/0022-4537.00239

Sipe, S., Johnson, C. D., & Fisher, D. K. (2009). University Students' Perceptions of Gender Discrimination in the Workplace: Reality Versus Fiction. *Journal of Education for Business*, *84*(6), 339–349. https://doi.org/10.3200/JOEB.84.6.339-349

Stanculescu, M., Marinescu, S., & Cheregi, G. (2009). MATLAB in Electrical Engineering. *Journal of Electrical and Electronics Engineering*, *2*(10).

United Nations Development Programme. (2023). *Gender social norms index (gsni)* [Accessed: March 22, 2025].

Wilkinson, A., & Barry, M. (2020). *The Future of Work and Employment*. Edward Elgar Publishing.

Yigit, Y., & Ali, Y. (2024). Generational Differences in Attitudes towards Gender Roles and Violence against Women. *International Journal of Caring Sciences*, *17*(1), 198–208.